13/PRTS JC07 Rec'd PCT/PTO 13 DEC 2001
PCT/KR01/00549
10/009792

ESCHERICHIA COLI STRAIN SECRETING HUMAN GRANULOCYTE COLONY STIMULATING FACTOR(G-CSF)

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to an *E. coli* producing and secreting human granulocyte-colony stimulating factor(hG-CSF), more specifically, to a recombinant plasmid constructed to express secretory hG-CSF in *E. coli*, an *E. coli* transformed with the said plasmid to secrete hG-CSF, and a process for preparing hG-CSF using the said transformed *E. coli*.

Description of the Prior Art

Colony stimulating factors(CSFs) known to be synthesized by various cell types such as mononuclear macrophages, T-lymphocytes, and fibroblasts are found in the various parts of normal human body. CSFs are classified into three major categories, granulocyte-colony stimulating factor(G-CSF), macrophage-colony stimulating factor(M-CSF) and granulocyte/macrophage-colony stimulating factor(GM-CSF), among them, G-CSF is an essential protein in manufacturing blood cells via promoting proliferation and differentiation of hemopoietic stem cells, and facilitates increase in numbers of granulocytes, especially, neutrophils which play an important role in the protection of the body from the infection. Chemotherapies widely used to treat growing tumors not only inhibit the growth of tumors but also inhibit the production of neutrophils, giving rise to severe side effects due to the diminished protecting function of neutrophils. Administration of G-CSF to the patients under such chemotherapies is known to be an effective way of treatment and prevention of the

infectious diseases by means of facilitating the increase in neurophil numbers.

In 1986, the hG-CSF gene was isolated from a human squamous carcinoma cell line CHU-II, its nucleotide sequence was first determined and expressed in COS cells by Nagata et al.(see: Nagata et al., *Nature* 319: 415 (1986)). The hG-CSF is a glycoprotein comprising 30 signal peptides which consists of 174 amino acid residues. The hG-CSF includes 5 cysteine residues of which 4 cysteines form two disulfide bonds, between Cys-36 and Cys-64, and between Cys-64 and Cys-74, which serve for folding of the expressed protein and its activity(see: Hill et al., *Proc. Natl. Acad. Sci.*, USA, 90:5167-5171(1993)). The hG-CSF dose not have the consensus sequence(Asn-X-Thr/Ser) for N-glycosylation, but O-glycosylation occurs at Thr-133. However, the recombinant G-CSF produced in *E. coli* is known to have almost the same biological activities as natural G-CSF, which means glycosylation is unnecessary for the G-CSF activity.

With the recent progress in recombinant DNA technology, G-CSF can be produced in bacteria, plant cells and animal cells, and some results previously reported are described here: Souza et al. isolated a cDNA from the human bladder cancer cell line 5637, determined its sequence and reported its expression in *E. coli*(see: Souza et al., *Science*, 232:61(1986)). Moreover, researches on production of hG-CSF in *E. coli*, plant cells and animal cells, and on construction and production of hG-CSF derivatives have been reported. However, technologies known to date for the production of hG-CSF in *E. coli* have many disadvantages in terms of protein yield or production cost since the hG-CSF is produced in cytoplasm in the form of insoluble inclusion body, which requires subsequent solubilization and renaturation to obtain biologically active form of hG-CSF protein. Although small quantity of soluble hG-CSF can be isolated directly, such method still have limitations in a sense that the active fraction of hG-CSF protein has to be

isolated from the pool of enormous amounts of *E. coli* proteins.

In general, proteins secreted to the periplasm of *E. coli* carry signal sequence, which is found in all proteins transportable out of the cytoplasm, and cleaved off by signal peptidase in the periplasm. The signal sequence is essential in secreting proteins in *E. coli*. Therefore, recombinant proteins originally not encoded by *E. coli* genes can be secreted into the periplasm or to the extracellular broth by joining known signal sequence(OmpA, OmpF, PelB, PhoA, SpA, etc.), as it is or with slight modifications, to the N-terminus of gene coding an exogenous protein.

The method for production of hG-CSF by secretion into the periplasmic space has following advantages over the conventional method by producing in cytoplasm described above: first, it is easy to isolate and purify the recombinant proteins, with high purity, in periplasm than in cytoplasm, since there are fewer proteins in periplasm than in cytoplasm(see: Nossal, N.G. et al., *J. Biol. Chem.*, 241:3055-3062(1966)); secondly, recombinant proteins secreted into periplasm are segregated from the cytoplasm where the most proteases are found, obtaining high yield in production of recombinant protein by avoiding degradation of the protein by proteases present in cytoplasm(see: Meerman and Georgiou, *Ann. N.Y. Acad. Sci.*, 721:292-302(1994)); thirdly, the bacterial periplasm is more oxidizing environment than cytoplasm, conducting disulfide bond formation and correct folding of polypeptide easily, thus, the formation of insoluble aggregates is avoidable(see: Hockney, *TIBTECH*, 12:456-463(1994)).

Having such advantages, the method resulting in secretion of recombinant proteins into the periplasm has been employed for production of hG-CSF in *E. coli* and reported as follows: Perez-Perez et al. have tried to get secreted form of hG-CSF, employing OmpA which is one of the signal sequences in *E. coli*, without success. To solve that

problem, they employed the system for coexpression of two molecular chaperones, DnaK and DnaJ, and merely obtain a small quantity of secreted hG-CSF(see: Perez-Perez et al., *Biochem. Biophys. Res. Commun.*, 210:524-529(1995));and, Chung et al. have tried to obtain secreted form of hG-CSF employing another signal sequence, PelB, again without success, but hG-CSF was accumulated in the form of insoluble inclusion body in cytoplasm(see: Chung et al., *J. Fermen. Bioengin.*, 85:443-446(1998)).

In view of above situation, there is a continuing need to develop the technique for facilitating secretion of hG-CSF at a substantial level into the periplasm through rather simple process which does not require solubilization of insoluble inclusion body and subsequent refolding.

SUMMARY OF THE INVENTION

The present inventors have made an effort to develop the technique for facilitating secretion of hG-CSF into the periplasm, thus, they have found that *E. coli*, transformed with the recombinant plasmid containing nucleotide sequence coding for oligopeptide consisting of 13 amino acids, *Bascillus*-derived endoxylanase signal sequence and 6 histidine residues, can secrete oligopeptide/hG-CSF fusion protein efficiently so that hG-CSF protein can be produced using the said transformed *E. coli*.

The first object of the present invention is, therefore, to provide the recombinant plasmid constructed for *E. coli* to express and secrete hG-CSF fusion protein therefrom.

The second object of the invention is to provide transformed *E. coli* with the said plasmid.

The third of the invention is to provide a process for producing hG-CSF protein employing the said

microorganism.

BRIEF DESCRIPTION OF THE DRAWINGS

The above, the other objects and features of the invention will become apparent from the following descriptions given in conjunction with the accompanying drawings, in which:

Figure 1 is the nucleotide sequence of hG-CSF gene inserted into the plasmid p19CSF.

Figure 2 shows a construction scheme and a genetic map for the plasmid p19CSFm.

Figure 3 is the nucleotide sequence and the deduced amino acid sequence of hG-CSF gene inserted into the plasmid p19CSFm.

Figure 4 shows a construction scheme and a genetic map for the plasmid pEDCSFm.

Figure 5 is the nucleotide sequence and the deduced amino acid sequence of hG-CSF gene inserted into the plasmid pEDCSFm.

Figure 6 shows a construction scheme and a genetic map for the plasmid pEDCSFmII.

Figure 7 is the nucleotide sequence and the deduced amino acid sequence of the N-terminal portion of hG-CSF gene inserted into the plasmid pEDCSFmII.

Figure 8 shows a construction scheme and a genetic map for the plasmid pTrcSCSFmII.

Figure 9 is the nucleotide sequence and the deduced amino acid sequence of the N-terminal portion of hG-CSF gene inserted into the plasmid pTrcSCSFmII.

Figure 10 shows a construction scheme and a genetic map for the plasmid pTrcKCSFmII.

Figure 11 shows a construction scheme and a genetic map for the plasmid pTHSCSFmII.

Figure 12 is the nucleotide sequence and the deduced amino acid sequence of the N-terminal portion of hG-CSF gene inserted into the plasmid pTHSCSFmII.

Figure 13 shows a construction scheme and a genetic map for the plasmid pTHKCSFmII.

DETAILED DESCRIPTION OF THE INVENTION

The recombinant plasmid of the invention designed for secreting hG-CSF fusion protein efficiently from *E. coli* contains: a kanamycin resistance gene, an endoxylanase signal sequence, a nucleotide sequence coding for an oligopeptide consisting of 13 amino acid residues including 6 consecutive histidine residues, a modified hG-CSF gene and Trc promoter. The recombinant *E. coli* of the invention secreting hG-CSF protein efficiently is prepared by transformation of *E. coli* with the said recombinant plasmid. The hG-CSF protein can be prepared with high purity by isolating the hG-CSF fusion protein from the protein pool obtained from the transformed *E. coli* using a Ni-column and subsequent protease treatment.

The present invention is illustrated in more detail as followings.

First, the recombinant plasmid vector pTHKCSFmII, containing cDNA encoding hG-CSF protein, endoxylanase signal sequence for hG-CSF protein secretion, Trc promoter which is a strong inducible promoter, kanamycin resistance gene encoding non-secretory protein, and a nucleotide sequence encoding an oligopeptide consisting of 13 amino acid residues including 6 consecutive histidine residues between endoxylanase signal sequence and hG-CSF polypeptide, is constructed for preparing transformed *E. coli* which expresses and secretes hG-CSF fusion protein into the periplasm: wherein, cDNA encoding hG-CSF protein is

obtained by joining 3rd exon-deleted hG-CSF cDNA obtained from a human breast carcinoma cDNA library to the synthesized 3rd exon, *Bacillus* sp.-derived endoxylanase signal sequence is employed for facilitating secretion of hG-CSF protein from *E. coli*, and kanamycin resistance gene is employed as a selection marker. Furthermore, to prevent cell lysis occurred during the secretion of hG-CSF protein in *E. coli*, a nucleotide sequence encoding an oligopeptide consisting of 13 amino acid residues including 6 consecutive histidine residues was included between DNA of endoxylanase signal sequence and hG-CSF gene, and the amino acid sequence of the oligopeptide is N'-Ala-Gly-Pro-His-His-His-His-His-Ile-Glu-Gly-Arg-C'(SEQ ID NO: 1).

Subsequently, recombinant *E. coli* secreting hG-CSF fusion protein is prepared by introducing the plasmid pTHKCSFmII constructed above into *E. coli*: wherein, *E. coli* strains BL21(DE3), HB101, MC4100, W3110 and XL1-Blue, preferably MC4100, can be used. *E. coli* MC4100 transformed with the recombinant plasmid pTHKCSFmII was named *E. coli* MC4100/pTHKCSFmII(*Escherichia coli* MC4100/pTHKCSFmII), which was deposited with the Korean Collection for Type Cultures(KCTC) affiliated to Korea Research Institute of Bioscience and Biotechnology(KRIBB), an international depository authority, under accession(deposition) No. KCTC 0754BP on Mar. 13, 2000.

Finally, the complete hG-CSF protein is prepared from the protein pool obtained from the transformed *E. coli* using a Ni-column and a protease: the said *E. coli* secretes hG-CSF fusion protein of which N-terminus is linked by an oligopeptide consisting of 13 amino acid residues including 6 consecutive histidine residues. The secreted hG-CSF fusion protein is isolated using a Ni-column to which 6 consecutive histidine residues present in the oligopeptide of the fusion protein can bind, and then the complete hG-CSF protein can be prepared from the hG-CSF fusion protein isolated above by treating a protease to get rid of the oligopeptide. Since the hG-CSF protein has to be non-

susceptible to the protease employed, the C-terminal sequence of the oligopeptide should be selected to be cleaved off by the protease of which recognition sequences are not present in the hG-CSF protein. As an example, in the present invention, the C-terminal amino acid sequence of the oligopeptide was selected to be Ile-Glu-Gly-Arg(see: SEQ ID NO: 1), which is recognized and cleaved by Factor Xa, a protease not having recognition sequences in hG-CSF protein.

The present invention is further illustrated in the following examples, which should not be taken to limit the scope of the invention.

Example 1: Preparation of cDNA for human granulocyte-colony stimulating factor(hG-CSF)

To isolate hG-CSF gene, hG-CSF cDNA was prepared by PCR amplification of a human breast carcinoma cDNA library. Nucleotide sequence of the hG-CSF gene was retrieved in the GenBank, and then, primer 1: 5'-GCGAATTCATGGCTGGACCTGCCACCCAG-3'(SEQ ID NO: 2) and primer 2: 5'-GCGGATCCTTATTAGGGCTGGGCAAGGTGGCG-3'(SEQ ID NO: 3) were synthesized, respectively. For the convenience of cloning of PCR product, *Eco*RI and *Bam*HI restriction sites were introduced at primer 1 and primer 2, respectively. Then, PCR was performed by, employing High Fidelity PCR System(Boehringer Mannheim Co., Germany), under following condition: one cycle of denaturation at 94℃ for 7 min; 30 cycles of denaturation at 94℃ for 1 min, annealing at 50℃ for 1 min, and extension at 72℃ for 1 min; and, one cycle of extension at 72℃ for 7 min. DNA molecules obtained by the PCR were subjected to agarose gel electrophoresis to isolate DNA fragment of approximately 530bp, which was subsequently digested with *Eco*RI and *Bam*HI to obtain DNA fragment concerned.

In order to clone the DNA fragment obtained above,

the said DNA fragment was joined into the plasmid pUC19(see: Yanisch-Perron et al., Gene, 33:109-119(1985)) digested with *Eco*RI and *Bam*HI using T4 DNA ligase, and then, the ligation product was introduced into *E. coli* XL1-Blue(*supE44 hsdR17 recA1 endA1 gyrA96 thi relA1 lacF'*(*proAB+ lacIqlacZΔ M15Tn*(tetr)))(see: Bullock, W. O. et al., *BioTechniques*, 5:376-378(1987)) by electroporation technique to obtain transformed *E. coli*. The transformants were selected on the LB agar medium containing ampicillin(50 μg/ℓ) and the recombinant plasmid p19CSF was obtained therefrom. The nucleotide sequence of the fragment in p19CSF was determined using automatic DNA sequencer(ABI Prism model 377, Perkin Elmer Co., U.S.A.). It has been found that there was 108bp deletion in the middle of the cloned fragment of the invention compare to the published nucleotide sequence of hG-CSF gene, meanwhile, the rest sequence of 508bp was identical to the published sequence of hG-CSF gene(see: Figure 1). Comparing with the sequence of genomic DNA of hG-CSF, the 108bp deletion in the fragment of the invention corresponded to the third exon among five exons of hG-CSF gene.

In order to obtain complete sequence of the hG-CSF gene, four primers, primer 3: 5'-TCCTCGGGGTGGCACAGCTTGTAGGTGGCACACAGCTTCTCCTGGAGCGC-3'(SEQ ID NO: 4), primer 4: 5'-GCTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCC-3'(SEQ ID NO: 5), primer 5: 5'-TGGCTGGGGCAGCTGCTCAGGGGAGCCCAGGGGATGCCCAGAGAGTGTC-3'(SEQ ID NO: 6), and primer 6: 5'-AGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCTGCTTGAGCCAA-3'(SEQ ID NO: 7) corresponding the third exon were prepared and PCR was carried out. Using primer 1 and the PCR product obtained above, p19CSF was subjected to PCR amplification, and with the primer 2, PCR was carried out in the same way as with primer 1. Using both PCR products, PCR was performed once again and the resulting product was digested with two restriction enzymes, *Eco*RI and *Bam*HI, and then the

fragments were cloned into the same site of pUC19 as the site used above. The transformed E. coli XL1-Blue were selected on the LB agar medium containing ampicillin(50 μg/ℓ) and the recombinant plasmid p19CSFm was obtained therefrom(see: Figure 2). The nucleotide sequence of the fragment in p19CSFm was determined and found to be identical to the sequence of the published hG-CSF gene(see: Figure 3).

Example 2: Construction of the recombinant plasmid, pEDCSFm

In order to get hG-CSF protein expressed from the hG-CSF gene obtained in Example 1, the mature hG-CSF gene was constructed by removing signal sequence from p19CSFm(see: Figure 5). To clone the mature hG-CSF gene, Primer 7: 5'-GCGAATTCATATGACCCCCCTGGGCCCTGCCAGC-3'(SEQ ID NO: 8) was prepared. Employing primer 2 and primer 7, p19CSFm was subjected to PCR amplification, and the amplified DNA molecules were digested with two restriction enzymes, *Nde*I and *Bam*HI. The resultant DNA fragment was joined into the T7 promoter-containing plasmid pET21c(Novagen Co., U.S.A.) digested with *Nde*I and *Bam*HI. After transformation of E. coli XL1-Blue with the ligated product by electroporation, the transformants were selected on the LB agar medium containing ampicillin(50 μg/ℓ) and the recombinant plasmid pEDCSFm was obtained therefrom(see: Figure 4).

For production of hG-CSF protein in *E. coli*, the recombinant plasmid pEDCSFm was introduced into *E. coli* BL21(DE3)(*F- ompT hsdSB(rB- mB-) gal dcm (DE3)* a prophage carrying the T7 RNA polymerase gene). The transformed *E. coli* harboring the recombinant plasmid was inoculated into 50ml of liquid LB medium in a 250ml flask and incubated at 37℃. When the cell concentration reached to OD600 of about 0.7, 1mM IPTG(isopropyl-β-thiogalactoside) was added to the culture to induce the expression of hG-CSF gene. After 4 hour induction, 1ml of the culture broth was aliquoted. The aliquoted broth was centrifuged at a speed

of 6,000rpm for 5min at 4℃ to collect cells, which were then washed, centrifuged again at a speed of 6,000rpm for 5min at 4℃, and then resuspended in 0.2ml of TE buffer. The mixture of 64 ㎕ aliquot of cell suspension and 16 ㎕ of sample buffer(Tris-HCl 60mM: 25% glycerol (v/v): 2% SDS(v/v): 2-mercaptoethanol 14.4mM: 0.1% bromophenol blue) was heated at 100℃ for 10min, and then subjected to SDS-PAGE(sodium-dodecyl sulfate polyacrylamide gel electrophoresis) in separating gel. After SDS-PAGE, the gel was stained in a staining solution(Coomassie brilliant blue R 0.25g/ℓ : methanol 40%(v/v), acetic acid 7%(v/v)) for 2 hours, and destained two times in destaining solution(methanol 40%(v/v), acetic acid 7%(v/v)) for 2 hours each time. The SDS-PAGE did not show the protein band corresponding to hG-CSF protein.

Example 3: Construction of the recombinant plasmid, pEDCSFmII

Based on the report by Devlin et al. that high G+C content in N-terminal portion of hG-CSF gene has inhibitory effect on transcription and translation of the gene(see: Devlin et al., *Gene*, 65:13-22 (1988)), primer 8: 5'-GCGAATTCATATGACTCCGTTAGGTCCAGCCAGC-3'(SEQ ID NO: 9) was prepared to obtain the gene which has lower G+C content such that the gene can be transcribed and translated efficiently in *E. coli*.

Plasmid p19CSFm was amplified by PCR using primer 8 and primer 2. The resulting PCR product was digested with *Nde*I and *Bam*HI, cloned into the same site of pET21c, and then transformed into *E. coli* XL1-Blue. The transformed *E. coli* XL1-Blue were selected on the LB agar medium containing ampicillin(50 ㎍/ℓ) and the recombinant plasmid pEDCSFmII was obtained therefrom(see: Figure 6). Comparing with the sequence of the fragment in pEDCSFm, the nucleotide sequence of the fragment in pEDCSFmII was found to be identical in N-terminal portion of the gene,

meanwhile, be different in 6 nucleotide residues overall, which give rise the changes in 5 codons(ACC CCC CTG GGC CCT → ACT CCG TTA GGT CCA).

Plasmid pEDCSFmII was transformed into *E. coli* BL21(DE3) and its expression was analyzed. Culturing of transformed *E. coli*, and production and analysis of total protein were carried out as described in Example 2, and fractionation of insoluble aggregate was performed as follows. Cells obtained by centrifugation of 1ml of culture broth were suspended in 0.5ml of TE buffer(Tris-HCl 10mM, EDTA 1mM, pH 8.0), and then centrifuged again at a speed of 3,000xg for 5min at 4℃ followed by resuspension in 0.2ml of TE buffer. Cells were disrupted by using a ultrasonicator(Branson Ultrsonics Co., U.S.A.) and then centrifuged at a speed of 10,000xg for 10min at 4℃. The supernatant was classified as soluble proteins and the pellet was dissolved in 0.2ml of TE buffer and classified as insoluble proteins. The content of hG-CSF protein expressed in *E. coli* BL21(DE3) transformed with plasmid pEDCSFmII was identified to be as much as about 40% of total protein, and, most of hG-CSF protein produced was found to be insoluble aggregates with no biological activity.

Example 4: Construction of the recombinant plasmid, pTrcSCSFmII

The signal sequence of endoxylanase derived from *Bacillus* sp. was used for the secretion of hG-CSF protein from *E. coli*. To join the signal sequence of endoxylanase to the N-terminal of hG-CSF polypeptide by PCR, primer 9: 5'-GGAATTCACATGTTTAAG TTTAAAAAGAAATTC-3'(SEQ ID NO: 10), primer 10: 5'-GGCTGGACCTAACGGAGTTGCAGAGGCGG-3'(SEQ ID NO: 11) and primer 11: 5'-GCAACCGCCTCTGCAACTCCGTTAGGTCCAGCC-3'(SEQ ID NO: 12) were prepared, respectively. From PCR performed using primers 9 and 10, and endoxylanase gene-

carrying plasmid pKJX4(see: Jeong et al., *Enzyme Microb Technol.* 22(7):599-605(1998)) as a target DNA, PCR product containing endoxylanase signal sequence was obtained, and from PCR performed using primers 2 and 11, PCR product containing hG-CSF gene was obtained. Again, PCR was performed using two PCR products obtained above, and primers 2 and 9 to obtain PCR product containing endoxylanase signal sequence-fused hG-CSF gene, which was then digested with two restriction enzymes, *Afl*III and *Bam*HI. Also, the plasmid pTrc99A(Pharmacia Biotech Co., U.S.A.) carrying a strong inducible promoter, trc promoter, was digested with *Nco*I and *Bam*HI, and then ligated with the above two PCR products. The ligation mixture was transformed into *E. coli* XL1-Blue. The transformants were selected on the LB agar medium containing ampicillin(50 μg/ℓ) and the recombinant plasmid pTrcSCSFmII was obtained therefrom(see: Figure 8). The nucleotide sequence of the N-terminal portion of hG-CSF gene was determined using an automatic DNA sequencer(ABI Prism model 377, Perkin Elmer Co., U.S.A.)(see: Figure 9).

In order to analyze the secretion of hG-CSF protein in various strains of *E. coli*, the recombinant plasmid pTrcSCSFmII was transformed into *E. coli* BL21(DE3), *E. coli* HB101(*F- hsdS20(rk-, mk-) recA13 ara-14 proA2 lacY1 ga1K2 rpsL20(strr) xyl1-5 mtl-1 supE44* λ -), *E. coli* MC4100(*F- araD139* Δ *(argF-lac)U169 rpsL150(strr) relA1 flbB5301 deoC1 ptsF25 rbsR*) and *E. coli* W3110(derived from *E. coli* K-12, λ -, F-, prototrophic). Each transformant was inoculated into 50ml aliquots of LB medium containing ampicillin(50 μg /ℓ), respectively, to perform the experiments for measuring expression level and protein analysis by the method described in Example 2. After induction with 1mM IPTG, however, the cells began to lyse abruptly, and within 2 hours most cells were lysed. The cultures were centrifuged to obtain transformed *E. coli* cells and proteins were analyzed using SDS-PAGE to find out no protein corresponding hG-CSF protein was obtained.

Example 5: Construction of the recombinant plasmid, pTrcKCSFmII

Considering the fact that ampicillin resistance gene(Apr) in plasmid pTrcSCSFmII is coding for β -lactamase which is secreted to the periplasm of *E. coli*, it may be conjectured that there was a competition between β -lactamase and hG-CSF protein for secretion, and such competition brought about cell lysis observed in Example 4. Based on such hypothesis, kanamycin resistance gene(Kmr) was employed instead of ampicillin resistance gene(Apr) as a selection marker since kanamycin resistance gene encodes the non-secreted protein, still can be used as a selection pressure. In order to substitute ampicillin resistance gene(Apr) in plasmid pTrcSCSFmII with kanamycin resistance gene(Kmr), PCR was performed using primer 12: 5'-GCGAATTCTTTAAAGCCACGTTGTGTCCTCAAA-3'(SEQ ID NO: 13) and peimer 13: 5'-GCGAATTCTTTAAATTAGAAAAACTCATCGAGCATC-3'(SEQ ID NO: 14), and plasmid pACYC177 as a target DNA. The resulting PCR product containing kanamycin resistance gene was digested with *Dra*I, and then ligated to the plasmid pTrcSCSFmII which was partially digested with *Dra*I to remove ampicillin resistance gene. The ligation product was transformed into *E. coli* XL1-Blue. The transformants were selected on LB agar medium containing kanamycin(25 μg/ℓ) and the recombinant plasmid pTrcKCSFmII was obtained therefrom(see: Figure 10).

To analyze the secretion of hG-CSF protein, the recombinant plasmid pTrcKCSFmII was transformed into *E. coli* BL21(DE3), *E. coli* HB101, *E. coli* MC4100 and *E. coli* W3110. Each transformant was inoculated into 50ml aliquots of LB medium containing kanamycin(25 μg/ℓ), respectively, to perform the experiments for measuring expression level and protein analysis by the method described in Example 2, and same result was obtained as in Example 4 wherein

pTrcSCSFmII was used. That is, with pTrcKCSFmII also, cells were lysed excessively upon induction with 1mM IPTG and SDS-PAGE analysis of recovered protein showed no protein corresponding to hG-CSF protein. From the above studies, it has been found that the substitution of ampicillin resistance gene with kanamycin resistance gene exert no effect on avoiding cell lysis, therefore, another genetic manipulation which can solve the intrinsic problem of cell lysis caused by secretion of hG-CSF protein is necessary.

Example 6: Construction of the recombinant plasmids, pTHSCSFmII and pTHKCSFmII

To facilitate *E. coli* to secrete hG-CSF protein without resulting in cell lysis, the strategy of inserting a small oligopeptide between endoxylanase signal sequence and hG-CSF protein was employed. For this purpose, PCR was performed using primer 14: 5'-CACCATCACCATATCGAAGGCCGTACTCCGTTAGGTCCA-3'(SEQ ID NO: 15) and primer 15: 5'-GATATGGTGATGGTGATGGTGCGGGCCAGCTGCAGAGGCGG-3'(SEQ ID NO: 16), and pEDSCSFmII as a target DNA. First, PCR with primers 14 and 2, and PCR with primers 15 and 9 were performed, respectively. Then, the two PCR products were mixed to be subjected to another PCR using primers 2 and 9, and the resulting product was digested with *Afl*III and *Bam*HI to join into the *Nco*I/*Bam*HI site of plasmid pTrc99A. After introducing the ligation product into *E. coli* XL1-Blue, the transformants were selected on a LB agar medium containing ampicillin(50 $\mu g/\ell$), and the recombinant plasmid pTHSCSFmII was obtained therefrom(see: Figure 11). Figure 12 shows the nucleotide sequence and deduced amino acid sequence of the inserted DNA in the plasmid pTHSCSFmII, wherein the italic indicates the endoxylanase signal sequence and the bold indicates inserted oligopeptide sequence.

As similarly as in Example 5, in order to substitute

ampicillin resistance gene(Apr) with kanamycin resistance gene(Kmr), the plasmid pTHSCSFmII was partial digested with DraI to remove ampicillin resistance gene(Apr) and then ligated with DNA fragment carrying kanamycin resistance gene obtained by digestion of pTrcKCSFmII with DraI. The ligation product was transformed into *E. coli* XL1-Blue. The transformants were selected on LB agar medium containing kanamycin(50 $\mu g/\ell$) and the recombinant plasmid pTHKCSFmII was obtained therefrom(see: Figure 13).

Example 7: Secretion of hG-CSF fusion protein from the transformed E. coli carrying the recombinant plasmid pTHKCSFmII

To analyze the expression and secretion of hG-CSF fusion protein from pTHKCSFmII, the recombinant plasmid pTHKCSFmII was transformed into *E. coli* BL21(DE3), *E. coli* HB101, *E. coli* MC4100 and *E. coli* W3110, respectively. Each transformant was cultured in 50ml aliquots of LB medium under temperatures of 37℃ and 30℃, respectively, followed by analyses of secretion of the hG-CSF fusion protein as described in Example 2. After 4 hour induction of hG-CSF gene expression with 1mM IPTG, 1ml aliquots of culture broth were centrifuged to obtain the cells, which were then analyzed by SDS-PAGE for the expression of hG-CSF fusion protein to show the secretion of hG-CSF fusion protein from all the transformed *E. coli*. The contents and the secretion efficiencies of hG-CSF fusion protein secreted from the each transformed *E. coli* are shown in Table 1.

Table 1: Comparison of production and secretion of hG-CSF in recombinant *E. coli*

Host Cell	Content of hG-CSF(%) in Total Protein		Secretion of hG-CSF(%)	
Culture Time	30°C	37°C	30°C	37°C
E. coli BL21(DE3)	22.7	22.1	>98	>98
E. coli HB101	13.5	12.8	81	75

E. coli MC4100	22.1	20.8	>98	>98
E. coli W3110	10.5	10.0	77	66
E. coli XL1-Blue	9.4	8.8	56	51

To examine if the secreted hG-CSF protein has been processed correctly, i.e., if the signal sequence has been removed correctly, hG-CSF fusion protein was isolated from the gel and its N-terminal amino acid sequence was determined to be N'-Ala-Gly-Pro-His-His-His-His-His-His-Ile-Glu-Gly-Arg-Thr-C', which is in agreement with the deduced amino acid sequence of N-terminal portion of hG-CSF fusion protein, indicating that the hG-CSF fusion protein was successfully secreted from *E. coli*. Of two temperature conditions, all the transformants showed higher secretion efficiency and higher hG-CSF protein content at 30℃ than 37℃. Of 5 strains of transformed *E. coli*, BL21(DE3) and MC4100 showed the highest production yield. Thus, *E. coli* MC4100 transformed with the recombinant plasmid pTHKCSFmII was named *E. coli* MC4100/pTHKCSFmII, which was deposited with the Korean Collection for Type Cultures(KCTC) affiliated to Korea Research Institute of Bioscience and Biotechnology(KRIBB), an international depository authority, under accession(deposition) No. KCTC 0754BP on Mar. 13, 2000.

The hG-CSF fusion polypeptide expressed in the transformed *E. coli* MC4100/pTHKCSFmII was processed during the secretion process to the periplasm through innermembrane to remove endoxylanase secretion sequence, resulting in secretion of oligopeptide/hG-CSF fusion protein. The hG-CSF fusion protein can be isolated easily by using a nickel column due to 6 histidine residues in N-terminal portion of the fusion protein, and C-terminus of the oliopeptide in the fusion protein can be recognized by Factor Xa and the oligopeptide is cleaved off to give a complete hG-CSF protein.

As clearly illustrated and demonstrated above, the

present invention provides an *E. coli* producing and secreting human granulocyte-colony stimulating factor(hG-CSF), a plasmid vector therefor, and a process for producing complete hG-CSF protein with high purity from the protein pool secreted by the said microorganism. In accordance with the invention, the hG-CSF protein can be prepared with high purity through rather simple process facilitating secretion of large amount of hG-CSF fusion protein into the periplasm, which does not require complicated processes such as solubilization and subsequent refolding required for isolation of the hG-CSF protein produced in cytoplasm as insoluble inclusion bodies by conventional techniques, thus, the hG-CSF protein can be widely used as an active ingredient in the development of supplementary agents for anticancer therapy.

It will be understood that the above description is merely illustrative of the preferred embodiment and it is not intended to limit the scope of the invention to the particular forms set forth, but on the contrary, it is intended to cover such alteranatives, modifications and equivalents as may be included within the spirit and scope of the invention as defined by the claims.

INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13*bis*)

A. The indications made below relate to the deposited microorganism or other biological material referred to in description
On page 7 , lines 18-25 .

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures (KCTC)

Address of depositary institution (including postal code and country)

Korea Research Institute of Bioscience and Biotechnology(KRIBB)
#52, Oun-dong, Yusong-ku,
Taejon 305-333,
Republic of Korea

Date of deposit	Accession Number
Mar. 13, 2000	KCTC 0754BP

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information continues on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1998)